CL

INPUT SET: S1916.raw

> This Raw Listing contains only the General
> Information Section and up to the first 5 pages.

```
 1                          SEQUENCE LISTING
 2
 3   (1)     General Information:
 4
 5        (i) APPLICANT: Celeste, Anthony J.
 6                       Wozney, John
 7                       Rosen, Vicki A.
 8                       Wolfman, Neil
 9                       Thomsen, Gerald H.
10                       Melton, Douglas A.
11
12      (ii) TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
13
14     (iii) NUMBER OF SEQUENCES: 24
15
16      (iv) CORRESPONDENCE ADDRESS:
17            (A) ADDRESSEE: GENETICS INSTITUTE, INC.
18            (B) STREET: 87 CambridgePark Drive
19            (C) CITY: Cambridge
20            (D) STATE: Massachusetts
21            (E) COUNTRY: USA
22            (F) ZIP: 02140
23
24       (v) COMPUTER READABLE FORM:
25            (A) MEDIUM TYPE: Floppy disk
26            (B) COMPUTER: IBM PC compatible
27            (C) OPERATING SYSTEM: PC-DOS/MS-DOS
28            (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
29
30      (vi) CURRENT APPLICATION DATA:
31            (A) APPLICATION NUMBER: US
32            (B) FILING DATE:
33            (C) CLASSIFICATION:
34
35    (viii) ATTORNEY/AGENT INFORMATION:
36            (A) NAME: Lazar, Steven R.
37            (B) REGISTRATION NUMBER: 32,618
38            (C) REFERENCE/DOCKET NUMBER: 5202-CIP
39
40      (ix) TELECOMMUNICATION INFORMATION:
41            (A) TELEPHONE: 617 876-1170
42            (B) TELEFAX: 617 876-5851
43
44
45   (2) INFORMATION FOR SEQ ID NO:1:
46
```

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/217,780*

INPUT SET: S1916.raw

```
47          (i) SEQUENCE CHARACTERISTICS:
48               (A) LENGTH: 926 base pairs
49               (B) TYPE: nucleic acid
50               (C) STRANDEDNESS: single
51               (D) TOPOLOGY: linear
52
53         (ii) MOLECULE TYPE: DNA (genomic)
54
55         (vi) ORIGINAL SOURCE:
56               (A) ORGANISM: Homo sapiens
57
58        (vii) IMMEDIATE SOURCE:
59               (B) CLONE: v1-1
60
61         (ix) FEATURE:
62               (A) NAME/KEY: mat_peptide
63               (B) LOCATION: 571..882
64
65         (ix) FEATURE:
66               (A) NAME/KEY: CDS
67               (B) LOCATION: 1..882
68
69
70         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
71
72    GCG CGT AAT ACG ACT CAC TAT AGG GCG AAT TGG GTA CGG GGC CCA GGC          48
73    Ala Arg Asn Thr Thr His Tyr Arg Ala Asn Trp Val Arg Gly Pro Gly
74    -190                -185            -180                -175
75
76    AGC TGG ACT TCT CCG CCG TTG CTG CTG CTG TCC ACG TGC CCG GGC GCC          96
77    Ser Trp Thr Ser Pro Pro Leu Leu Leu Leu Ser Thr Cys Pro Gly Ala
78                    -170            -165                -160
79
80    GCC CGA GCG CCA CGC CTG CTG TAC TCG CGG GCA GCT GAG CCC CTA GTC          144
81    Ala Arg Ala Pro Arg Leu Leu Tyr Ser Arg Ala Ala Glu Pro Leu Val
82                -155            -150                -145
83
84    GGT CAG CGC TGG GAG GCG TTC GAC GTG GCG GAC GCC ATG AGG CGC CAC          192
85    Gly Gln Arg Trp Glu Ala Phe Asp Val Ala Asp Ala Met Arg Arg His
86            -140            -135            · -130
87
88    CGT CGT GAA CCG CGC CCC CCC CGC GCG TTC TGC CTC TTG CTG CGC GCA          240
89    Arg Arg Glu Pro Arg Pro Pro Arg Ala Phe Cys Leu Leu Leu Arg Ala
90          -125            -120            -115
91
92    GTG GCA GGC CCG GTG CCG AGC CCG TTG GCA CTG CGG CGA CTG GGC TTC          288
93    Val Ala Gly Pro Val Pro Ser Pro Leu Ala Leu Arg Arg Leu Gly Phe
94    -110            -105            -100                -95
95
96    GGC TGG CCG GGC GGA GGG GGC TCT GCG GCA GAG GAG CGC GCG GTG CTA          336
97    Gly Trp Pro Gly Gly Gly Gly Ser Ala Ala Glu Glu Arg Ala Val Leu
98                    -90            -85                -80
99
```

INPUT SET: S1916.raw

```
100    GTC GTC TCC TCC CGC ACG CAG AGG AAA GAG AGC TTA TTC CGG GAG ATC        384
101    Val Val Ser Ser Arg Thr Gln Arg Lys Glu Ser Leu Phe Arg Glu Ile
102                -75                 -70                 -65
103
104    CGC GCC CAG GCC CGC GCG CTC GGG GCC GCT CTG GCC TCA GAG CCG CTG        432
105    Arg Ala Gln Ala Arg Ala Leu Gly Ala Ala Leu Ala Ser Glu Pro Leu
106                -60                 -55                 -50
107
108    CCC GAC CCA GGA ACC GGC ACC GCG TCG CCA AGG GCA GTC ATT GGC GGC        480
109    Pro Asp Pro Gly Thr Gly Thr Ala Ser Pro Arg Ala Val Ile Gly Gly
110                -45                 -40                 -35
111
112    CGC AGA CGG AGG AGG ACG GCG TTG GCC GGG ACG CGG ACA GCG CAG GGC        528
113    Arg Arg Arg Arg Arg Thr Ala Leu Ala Gly Thr Arg Thr Ala Gln Gly
114    -30                 -25                 -20                 -15
115
116    AGC GGC GGG GGC GCG GGC CGG GGC CAC GGG CGC AGG GGC CGG AGC CGC        576
117    Ser Gly Gly Gly Ala Gly Arg Gly His Gly Arg Arg Gly Arg Ser Arg
118                -10                 -5                   1
119
120    TGC AGC CGC AAG CCG TTG CAC GTG GAC TTC AAG GAG CTC GGC TGG GAC        624
121    Cys Ser Arg Lys Pro Leu His Val Asp Phe Lys Glu Leu Gly Trp Asp
122                 5                  10                  15
123
124    GAC TGG ATC ATC GCG CCG CTG GAC TAC GAG GCG TAC CAC TGC GAG GGC        672
125    Asp Trp Ile Ile Ala Pro Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly
126                20                  25                  30
127
128    CTT TGC GAC TTC CCT TTG CGT TCG CAC CTC GAG CCC ACC AAC CAT GCC        720
129    Leu Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala
130     35                  40                  45                  50
131
132    ATC ATT CAG ACG CTG CTC AAC TCC ATG GCA CCA GAC GCG GCG CCG GCC        768
133    Ile Ile Gln Thr Leu Leu Asn Ser Met Ala Pro Asp Ala Ala Pro Ala
134                    55                  60                  65
135           '
136    TCC TGC TGT GTG CCA GCG CGC CTC AGC CCC ATC AGC ATC CTC TAC ATC        816
137    Ser Cys Cys Val Pro Ala Arg Leu Ser Pro Ile Ser Ile Leu Tyr Ile
138                70                  75                  80
139
140    GAC GCC GCC AAC AAC GTT GTC TAC AAG CAA TAC GAG GAC ATG GTG GTG        864
141    Asp Ala Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val
142                85                  90                  95
143
144    GAG GCC TGC GGC TGC AGG TAGCGCGCGG GCCGGGGAGG GGGCAGCCAC            912
145    Glu Ala Cys Gly Cys Arg
146        100
147
148    GCGGCCGAGG ATCC                                                      926
149
150
151    (2) INFORMATION FOR SEQ ID NO:2:
152
```

INPUT SET: S1916.raw

```
153              (i) SEQUENCE CHARACTERISTICS:
154                   (A) LENGTH: 294 amino acids
155                   (B) TYPE: amino acid
156                   (D) TOPOLOGY: linear
157
158          (ii) MOLECULE TYPE: protein
159
160          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
161
162   Ala Arg Asn Thr Thr His Tyr Arg Ala Asn Trp Val Arg Gly Pro Gly
163   -190                -185                -180                -175
164
165   Ser Trp Thr Ser Pro Pro Leu Leu Leu Ser Thr Cys Pro Gly Ala
166               -170                -165                -160
167
168   Ala Arg Ala Pro Arg Leu Leu Tyr Ser Arg Ala Ala Glu Pro Leu Val
169           -155                -150                -145
170
171   Gly Gln Arg Trp Glu Ala Phe Asp Val Ala Asp Ala Met Arg Arg His
172           -140                -135                -130
173
174   Arg Arg Glu Pro Arg Pro Pro Arg Ala Phe Cys Leu Leu Leu Arg Ala
175       -125                -120                -115
176
177   Val Ala Gly Pro Val Pro Ser Pro Leu Ala Leu Arg Arg Leu Gly Phe
178   -110                -105                -100                -95
179
180   Gly Trp Pro Gly Gly Gly Gly Ser Ala Ala Glu Glu Arg Ala Val Leu
181               -90                 -85                 -80
182
183   Val Val Ser Ser Arg Thr Gln Arg Lys Glu Ser Leu Phe Arg Glu Ile
184           -75                 -70                 -65
185
186   Arg Ala Gln Ala Arg Ala Leu Gly Ala Ala Leu Ala Ser Glu Pro Leu
187           -60                 -55                 -50
188
189   Pro Asp Pro Gly Thr Gly Thr Ala Ser Pro Arg Ala Val Ile Gly Gly
190       -45                 -40                 -35
191
192   Arg Arg Arg Arg Arg Thr Ala Leu Ala Gly Thr Arg Thr Ala Gln Gly
193   -30                 -25                 -20                 -15
194
195   Ser Gly Gly Gly Ala Gly Arg Gly His Gly Arg Arg Gly Arg Ser Arg
196               -10                 -5                  1
197
198   Cys Ser Arg Lys Pro Leu His Val Asp Phe Lys Glu Leu Gly Trp Asp
199           5                   10                  15
200
201   Asp Trp Ile Ile Ala Pro Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly
202       20                  25                  30
203
204   Leu Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala
205       35                  40                  45                  50
```

```
206
207    Ile Ile Gln Thr Leu Leu Asn Ser Met Ala Pro Asp Ala Ala Pro Ala
208                    55              60                65
209
210    Ser Cys Cys Val Pro Ala Arg Leu Ser Pro Ile Ser Ile Leu Tyr Ile
211                70                  75     ,          80
212
213    Asp Ala Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val
214            85                  90              95
215
216    Glu Ala Cys Gly Cys Arg
217        100
218
219    (2) INFORMATION FOR SEQ ID NO:3:
220
221        (i)   SEQUENCE CHARACTERISTICS:
222              (A) LENGTH: 1207 base pairs
223              (B) TYPE: nucleic acid
224              (C) STRANDEDNESS: single
225              (D) TOPOLOGY: linear
226
227        (ii)  MOLECULE TYPE: DNA (genomic)
228
229        (vi)  ORIGINAL SOURCE:
230              (A) ORGANISM: Homo sapiens
231
232       (vii)  IMMEDIATE SOURCE:
233              (B) CLONE: MP52
234
235        (ix)  FEATURE:
236              (A) NAME/KEY: CDS
237              (B) LOCATION: 845..1204
238
239
240        (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:3:
241
242    ACCGGGCGGC CCTGAACCCA AGCCAGGACA CCCTCCCCAA ACAAGGCAGG CTACAGCCCG    60
243
244    GACTGTGACC CCAAAAGGAC AGCTTCCCGG AGGCAAGGCA CCCCCAAAAG CAGGATCTGT   120
245
246    CCCCAGCTCC TTCCTGCTGA AGAAGGCCAG GGAGCCCGGG CCCCCACGAG AGCCCAAGGA   180
247
248    GCCGTTTCGC CCACCCCCCA TCACACCCCA CGAGTACATG CTCTCGCTGT ACAGGACGCT   240
249
250    GTCCGATGCT GACAGAAAGG GAGGCAACAG CAGCGTGAAG TTGGAGGCTG GCCTGGCCAA   300
251
252    CACCATCACC AGCTTTATTG ACAAAGGGCA AGATGACCGA GGTCCCGTGG TCAGGAAGCA   360
253
254    GAGGTACGTG TTTGACATTA GTGCCCTGGA GAAGGATGGG CTGCTGGGGG CCGAGCTCCG   420
255
256    GATCTTGCGG AAGAAGCCCT CGGACACGGC CAAGCCAGCG GCCCCCGGAG GCGGGCGGGC   480
257
258    TGCCCAGCTG AAGCTGTCCA GCTGCCCCAG CGGCCGGCAG CCGGCCTCCT TGCTGGATGT   540
```

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION   *US/08/217,780*

INPUT SET: S1916.raw

Line	Error	Original Text
31	Wrong application Serial Number	(A) APPLICATION NUMBER: US

ENTERED

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION   *US/08/217,780*

SEQUENCE MISSING ITEM REPORT
PATENT APPLICATION   *US/08/217,780*

INPUT SET: S1916.raw

APPLICATION NUMBER
FILING DATE
PRIOR APPLICATION DATA

SEQUENCE CORRECTION REPORT
PATENT APPLICATION *US/08/217,780*

INPUT SET: S1916.raw

Line	Original Text	Corrected Text